IST 691: Deep Learning in Practice

Homework 01

Instructions: Answer the following questions in no more than one page per question. In addition to the accuracy of your responses, the clarity, coherence, and concision of your writing are critical factors to earning full credit for this assignment. Cite any source you use outside of lecture notes. This includes the textbook. Reproducing or even paraphrasing an answer from a generative AI tool, such as ChatGPT, is not allowed on this assignment. Submit your responses in single Word or pdf document.

1. In traditional machine learning (non-NN methods), we use feature engineering to model complex relationships between observed variables (features) and a target variable (response). When using deep learning methods, should we design and incorporate feature engineering processes? Explain why or why not.

2. Explain in words or equations, why we should introduce nonlinearity in neural networks.

3. You are training a deep learning model to predict sentiment of Twitter posts — the model predicts whether a post is "happy" or "sad". Your model achieves 0.95 accuracy on the dataset you used to train the model. But when you take new posts from Twitter and use your model to predict the sentiment, the model performs much worse. What might have happened? What should you do to improve your model?

4. The MNIST dataset consists of images of dimension 28x28 pixels with one color channel (28x28x1), with each image corresponding to a label between 1 and 10. To build a classifier, we implement a multi-layer perceptron model with 3 hidden layers. The first two hidden layers have 100 perceptrons each, and the third hidden layer has 30 perceptrons. Calculate how many weights will be updated for each iteration of gradient descent. Show your work.

5. Answer the following questions based on a close reading of this article and possibly additional research (remember to cite your sources).

   https://www.nytimes.com/2023/06/28/technology/facial-recognition-shoplifters-britain.html

   a) What is the technology being discussed in this article? How does it relate to deep learning?
   b) What are some ethical concerns discussed in the article? Do you share these concerns? Why or why not?
   c) Are these concerns addressed by the company deploying this technology? How?
   d) What is your personal view of the way this technology is being used? Explain your reasoning.